

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>H04L</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 00/67412</b> <b>(43) International Publication Date:</b> 9 November 2000 (09.11.00)
<b>(21) International Application Number:</b> PCT/US00/11488 <b>(22) International Filing Date:</b> 28 April 2000 (28.04.00)  <b>(30) Priority Data:</b> 60/131,898      30 April 1999 (30.04.99)      US  <b>(71) Applicant:</b> DRYKEN TECHNOLOGIES [US/US]; 5731 Lyons View Drive, Suite 95, Knoxville, TN 37919 (US).  <b>(72) Inventors:</b> BLACK, Christopher, L.; 131 Northshore Drive, Anderson, TN 37705 (US). HINES, J., Wesley; 7654 Racing Run Road, Knoxville, TN 37920 (US).  <b>(74) Agent:</b> ROURK, Christopher, J.; Akin, Gump, Strauss, Hauer & Feld, L.L.P., 1700 Pacific Avenue, Suite 4100, Dallas, TX 75201-4675 (US).		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> METHOD AND SYSTEM FOR NONLINEAR STATE ESTIMATION  <b>(57) Abstract</b>  An NSET method and system for modeling the behavior of a visitor to an e-commerce location are disclosed. The method of one embodiment of this invention comprises the steps of: obtaining one or more visitor characteristic values; developing a model of the visitor's behavior according to a nonlinear state estimation technique (NSET); and estimating a set of visitor behavior characteristic to model the visitor's behavior using the developed model. The method further comprises predicting the visitor's future behavior at the e-commerce location based on the set of behavior characteristic values and known statistical behavior characteristic values. The method of this invention can further comprise measuring a set of actual behavior values for a visitor based on the visitor's actual behavior at an e-commerce location, and comparing the set of estimated behavior characteristic values and the set of actual behavior values with at least one similarity operator. The method of this invention can be implemented as a system of operational instructions that can be stored in a memory and executed by a processing module.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## METHOD AND SYSTEM FOR NONLINEAR STATE ESTIMATION

TECHNICAL FIELD OF THE INVENTION

5 This invention relates generally to methods and systems for process modeling and analysis. More particularly, the present invention relates to an improved nonlinear state estimation technique (NSET) to perform process modeling and analysis of, for example, buyer purchasing characteristics.

10

This application claims priority under 35 U.S.C. § 119(e) to provisional application number 60/131,898, filed April 30, 1999, by Black, et al., which is hereby fully incorporated by reference.

BACKGROUND OF THE INVENTION

Systems and methods for process modeling and analysis are well known in the art. For example, artificial neural networks, such as those disclosed in  
5 co-pending patent application having an attorney docket number of 103773-991110-1, entitled "An Improved Method and System For Training An Artificial Neural Network," having a filing date of March 31, 1999, Serial No. 09/282,392, and assigned to the same assignee as the  
10 present patent application, multi-variate state estimation techniques (MSET), as described by Singer, et al., in a paper entitled "Analytical Enhancements of Automotive Century System Reliability," dated June 1994, and multiple regression techniques,  
15 all are used for process modeling and analysis in varying applications. Similarly, prior art basic nonlinear state estimation techniques are also known for use in process modeling. Each of these prior art methods, however, suffers from limitations in terms of  
20 accuracy and modeling flexibility.

Artificial neural networks, for example, although suitable for modeling certain systems, require extensive training and are time-intensive, which makes them unsuitable for applications in which a system, and  
25 corresponding modeling of that system, must be done in near real time. An artificial neural network would thus be unsuitable, for example, to predict behavior in a e-commerce setting where the future behavior of a customer is desired to be known. Applying artificial

neural networks to model the behavior of each customer in such an application, in which new information (in the form of additional variables) becomes available as time evolves, is not possible, as means do not exist for rapid adjustment of the model of such a system to predict behavior. The iterative process required to train an artificial neural network is not conducive to modeling rapidly changing systems in which a rapid model adjustment is necessary once one or more new variables has become available.

MSETs and basic NSETs also face limitations in that they rely upon the inversion of data matrices (recognition matrices) that are sometimes singular (in which case inversion is impossible) or near-singular, in which case inversion is possible but end result prediction accuracy is negatively affected. Furthermore, MSETs have poor stability with respect to choice of data included in the prototype matrix, i.e., the inclusion/exclusion of any particular single data point in the prototype matrix can unduly affect prediction results. This is actually a result of co-linearities among the prototypical data points.

Lastly, the distance/similarity function typically chosen for use in MSETs is selected based upon its tendency to produce relatively well-conditioned (when compared to other distance/similarity functions) recognition matrices. Condition is an inverse measure of singularity (i.e., well-conditioned implies non-singular, which is good, while poorly-conditioned

implies near-singular, which is bad). Such a distance/similarity function, while generally providing better-conditioned recognition matrices, is not optimal in terms of accuracy and modeling flexibility.

SUMMARY OF THE INVENTION

Therefore, a need exists for a method and system for nonlinear state estimation for process modeling and analysis that can achieve improved matrix conditioning, improved model stability and improved prediction accuracy. Such a method and system would use regularization principles in the inversion of prototype matrices, so as to avoid prior art problems encountered in the inversion of troublesome singular or near-singular prototype matrices.

An ever further need exists for an NSET with improved stability in the selection of datasets included in the prototype matrix, that can reduce or eliminate co-linearities among the prototypical data points.

A still further need exists for an NSET method and system that uses a distance/similarity function optimized to provide greater accuracy and modeling flexibility.

Even further, a need exists for an NSET method and system that can use a broader spectrum of distance functions, thereby allowing for the development of more flexible system models that can be optimized to a particular data set or system to be modeled.

In accordance with the present invention, a method and system for nonlinear state estimation is provided that essentially eliminates or reduces disadvantages and problems associated with previously developed systems and methods for process modeling and analysis,

including the problems of non-optimal prototype matrix selection, instabilities with respect to choice of data, and non-optimal distance/similarity functions resulting in reduced accuracy and modeling flexibility.

5        More specifically, the present invention provides, in one embodiment, an NSET method and system for modeling the behavior of a visitor to an e-commerce location. The method of this embodiment comprises the steps of: obtaining one or more visitor characteristic  
10        values; developing a model of the visitor's behavior according to a nonlinear state estimation technique (NSET); and estimating a set of visitor behavior characteristic to model the visitor's behavior using the developed model. The method of this embodiment  
15        further comprises the step of predicting the visitor's future behavior at the e-commerce location based on the set of behavior characteristic values and known statistical behavior characteristic values.

20        The method of this invention can further comprise determining a set of residuals between the set of estimated behavior characteristic values and a set of actual behavior values and statistically monitoring the set of residuals. The NSET's parameters (the choice of similarity/distance function, and the number of  
25        prototypical datapoints) can then be adjusted to compensate for the residuals.

      The method of this invention can further likewise comprise determining residuals between the estimated current and future behavior characteristic values and a



set of desired behavior values. The e-commerce location can then be adjusted to compensate for the residuals. The adjusting of the e-commerce location can comprise adjusting goods, services and/or  
5 advertising provided at the e-commerce location.

Developing the model of a visitor's behavior comprises selecting a representative sample data set, based on the visitor characteristic values, from a set of statistical characteristic data within a historical  
10 database. The visitor characteristic values can comprise visitor demographic information and visitor purchase habits.

The method of this invention can further comprise measuring a set of actual behavior values for a visitor  
15 based on the visitor's actual behavior at an e-commerce location, and comparing the set of estimated behavior characteristic values and the set of actual behavior values with at least one similarity operator. The method of this invention can be implemented as a system  
20 of operational instructions that can be stored in a memory and executed by a processing module.

An important technical advantage of the method and system for nonlinear state estimation for process  
modeling and analysis of this invention is that it can  
25 achieve improved matrix conditioning, improved model stability and improved prediction accuracy using principles of regularization in the inversion of prototype matrices.

Another technical advantage of the NSET of this invention is improved stability in the selection of datasets included in the prototype matrix that can reduce or eliminate co-linearities among the prototypical data points.

Still another technical advantage of the NSET method and system of this invention is the use of a distance/similarity function optimized to provide greater accuracy and modeling flexibility.

Yet another technical advantage of the NSET method and system of the present invention is the ability to use a broader spectrum of distance functions that allow for the development of more flexible system models that can be optimized to a particular data set or system to be modeled.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete understanding of the present invention and the advantages thereof may be acquired by referring to the following description, taken in conjunction with the accompanying drawings in which like reference numbers indicate like features and wherein:

FIGURES 1 A, B and C illustrate three different embodiments of the NSET method and system of this invention can be used to perform process modeling of an instrument calibration verification system ;

FIGURE 2 illustrates the NSET Test Prediction Error using RMPE Operator and Varying Minkowski Parameter;

FIGURE 3 illustrates Estimated and Measured Cooling Tower Inlet Temperature (DIST - no scaling);

FIGURE 4 illustrates the Drifted Measurement and Estimate of Cooling Tower Inlet Temperature (DIST - no scaling);

FIGURE 5 illustrates the Drifted Measurement and NSET Estimate;

FIGURE 6 illustrates the SPRT Drift Diagnosis from NSET (DIST - no scaling) ISCV System;

FIGURE 7 illustrates the Estimated and Measured Outlet Temperature 1 (SERVO) (DIST - no scaling);

FIGURE 8 illustrates the Drifted Measurement and Estimate of Outlet Temperature 1 (SERVO) (DIST - no scaling);

FIGURE 9 illustrates the Drifted Measurement and NSET Estimate (Drift commencing at Sample 251);

FIGURE 10 illustrates the SPRT Drift Diagnosis from NSET (DIST - no scaling) ISCV System;

5       FIGURE 11 illustrates the Cooling Tower Inlet Temperature Measurement and CMLCC NSET Estimate; and

FIGURE 12 illustrates the Cooling Tower Inlet Temperature Measurement (Drifted) and CMLCC NSET Estimate.

DETAILED DESCRIPTION OF THE INVENTION

Preferred embodiments of the present invention are illustrated in the FIGURES, like numerals being used to refer to like and corresponding parts of various  
5 drawings.

A nonlinear state estimation technique (NSET) has been developed to perform process modeling and analysis. One embodiment of the NSET of this invention can be used to model a sensor and associated instrument  
10 channel calibration verification system. The model estimates the true process values, as functional sensors would provide them. The residuals between these estimates and the actual measurements (from sensors of unknown condition) can then be monitored  
15 using the sequential probability ratio test, a statistical decision method.

Still another embodiment of the NSET of this invention can be used in an electronic commerce (e-commerce) setting to model the behavior of human  
20 visitors to a web-site. For example, based on a visitor's demographic information or prior purchase history, future purchase activity by the customer can be predicted. The NSET of this invention can very rapidly adjust prediction models to account for any new  
25 piece of information relating to a visitor since that visitor's last visit to the e-commerce site.

The NSET method and system of this invention can be used to model many other systems than the illustrative embodiments discussed above. In

particular, the NSET of this invention can be used in the various applications discussed in copending patent application having an attorney docket number of 103773-991110-1, entitled "An Improved Method and System For  
5 Training An Artificial Neural Network," having a filing date of March 31, 1999, Serial No. 09/282,392, and assigned to the same assignee as the present patent application, hereby incorporated by reference in its entirety. The NSET of this invention can thus serve as  
10 a general technique for predictive and estimative modeling applicable to a variety of systems and applications.

The NSET of this invention is a generalization of the multivariate state estimation technique (MSET)  
15 described by Singer, et al. The MSET itself is an extension of the least-squares minimization of the multiple regression equation, incorporating proprietary comparison operators. The theoretical introduction to these estimation techniques is provided, and the NSET  
20 is demonstrated with several different comparison (either similarity or distance) operators, below. The estimation performance of the NSET for each of these operators is also evaluated below, resulting in a recommended operator for general process modeling  
25 purposes. The NSET is demonstrated, in one embodiment, as the modeling engine for a calibration verification system.

The NSET, MSET, and the multiple regression solution all utilize similarity operators to compare

new measurements to a set of prototypical measurements or states. This comparison process generates a weight vector that is used to calculate a weighted sum of the prototype vectors. The sum is weighted to provide an estimate of the true process values.

For the similarity function, multiple regression utilizes the dot product, the MSET uses one proprietary similarity/distance operators, and the NSET of this invention can incorporate one of several possible operators, many of which are presented, demonstrated and evaluated below.

The NSET of this invention functions as an regressive model, reproducing an estimate of as set of variables based upon the set of measured signals that are provided as inputs to the model. The training is single-pass (i.e., is not an iterative process), and consists of little more than the operations involved in a single matrix multiplication and an inversion or decomposition. Data selection plays an important role, as the number of operations (and processor time) required per recall is proportional to the product of the number of prototype measurements and the dimensionality of the measurements. Therefore, as a function of the number of signals to be monitored, and the data availability rate, there is an upper limit upon the number of patterns that may be included in the prototype measurement matrix. As the purpose of the prototype matrix is to compactly represent the entire

dynamic range of previously observed system states, the patterns that are included should be carefully chosen.

As an example of the multiple regression technique improved upon by the NSET of this invention, let A,  
5 referred to as the prototype matrix, represent a matrix assembled from selected column-wise measurement vectors, and let w, represent a vector of weights for averaging A to provide the estimated state y' as follows:

10 
$$\underline{y'} = \underline{A} \cdot \underline{w} \quad (\text{Eqn. 1})$$

Alternatively, variables other than those making up the 'state vector' may be predicted/estimated by augmenting or replacing the prototype matrix included in equation 1, such that it now contains additional  
15 variables. I.e., the weight vector is determined via the state vectors in the prototype matrix and their comparison to the new state vector, but is used to multiply other variables besides the ones included in the state vector.

20 The column-wise measurement vectors which make up the prototype matrix A are usually selected by a clustering data analysis technique. This selection is carefully performed to provide a compact, yet representative, subset of a large database of  
25 measurements spanning the full dynamic range of the system of interest. For example, in an e-commerce setting, the management vectors might be specific attributes from a historical database of visitor attributes that are shared by a current visitor.



An example of a prototype matrix, constructed from the vertices of a unit cube is given:

$$\underline{A} = [\underline{x}(1) \ \underline{x}(2) \ \dots \ \underline{x}(n)] = E.g., \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \quad (\text{Eqn. 2})$$

5

If  $\underline{\varepsilon}$  represents the difference between an observed state  $\underline{y}$  and the estimated state  $\underline{y}'$ , then the following relations may be constructed:

$$\underline{\varepsilon} = \underline{y} - \underline{y}' = \underline{y} - \underline{A} \cdot \underline{w} \quad (\text{Eqn. 3})$$

10

The least squares solution to the minimization of  $\underline{\varepsilon}$  yields the following expression for  $\underline{w}$ , (where the left hand factor of the matrix product is known as the recognition matrix):

$$\underline{w} = (\underline{A}^T \cdot \underline{A})^{-1} \cdot (\underline{A}^T \cdot \underline{y}) \quad (\text{Eqn. 4})$$

15

A chief liability of this method is that linear interrelationships between state vectors in  $\underline{A}$  result in conditioning difficulties associated with the inversion of the recognition matrix. This shortcoming is avoided by the NSET of this invention by applying nonlinear operators, as shown later, in lieu of the matrix multiplication. These operators generally result in better conditioned recognition matrices and more meaningful inverses of the recognition matrices.

20

25

The conditioning difficulty may also be ameliorated by solving the normal equations for the weight vector, as shown in Equation 5 below. Solving

the normal equations involves a decomposition of the recognition matrix and then elimination or back-substitution.

$$(\underline{A}^T \cdot \underline{A}) \cdot \underline{w} = \underline{A}^T \cdot \underline{y} \quad (\text{Eqn. 5})$$

5 Although the decomposition of the recognition matrix must only be performed once, up-front in a learning phase, the solution of the normal equations requires more operations per vector (than when simply using the matrix inverse) in the deployment or recall phase.

10 Unfortunately, the normal equations are not themselves immune to numerical instabilities. This is because zero or near zero pivot elements, due to correlation among prototype measurements, may be encountered during solution (elimination) of the normal  
15 equations. Zero and near-zero pivot elements result in, respectively, no solution, or a solution in which some of the fitted parameters (weights) have large magnitudes that delicately cancel one another out. Singular value decomposition (SVD) is a particularly  
20 stable (with regard to rank deficiencies) method of matrix decomposition for finding the pseudoinverse of a matrix. The SVD algorithm assigns small magnitude weights to prototype vectors whose combinations are irrelevant to the fit. This results in a fit that  
25 handles both overdetermined-ness (minimizes squared error) and underdetermined-ness (minimizes effect of correlation among prototypes), and as a result is numerically very dependable.

The NSET of this invention can use, as shown below, the SVD algorithm to provide a pseudo-inverse of the recognition matrix, which is then employed in the manner of the true inverse of equation 4. This pseudo-inversion is performed in a training phase prior to deployment of the model, and is considered analogous to the up-front training of an artificial neural network (ANN) model.

Another drawback (also necessitating the use of different similarity operators) to the least squares solution of the multiple regression equation is that perturbations outside the range observed in the prototype matrix in a single element of the observation vector result in large estimation errors. Whether or not this quality is truly a disadvantage is not clear, as a conservative model might be expected to fail dramatically, as opposed to giving reasonable looking, but inaccurate, results. In other words, it is questionable whether any propensity to always supply a reasonable appearing, though possibly incorrect result should be considered a positive trait. This is an issue implicit in any empirically derived modeling application in which the general form of the relationship is not known or presumed. Extrapolation features of the NSET (or of data-derived models in general) are outside the scope of this invention.

The NSET of this invention extends the multiple regression equations. This NSET of this invention (as

does MSET) springs forth from the multiple regression result as follows:

$$\underline{w} = (\underline{A}^T \oplus \underline{A})^{-1} (\underline{A}^T \oplus \underline{y}) \quad (\text{Eqn. 6})$$

or in normal equation form,

$$(\underline{A}^T \oplus \underline{A}) \underline{w} = \underline{A}^T \oplus \underline{y}. \quad (\text{Eqn. 7})$$

The  $\oplus$  symbol above represents any appropriate similarity or difference operator applied upon matrices. The definitions of a few of the many possible candidate operators that can be used with the NSET of this invention and provided below are all scalar-valued functions on vectors. The vector functions are easily extended to the required matrix functions by selecting vectors from the two input matrices and positions in the output matrix in exactly the familiar row and column manner used in performing individual vector dot products in matrix multiplication. Some of the possible candidates for distance/similarity operators include:

1. Bernoulli difference (BDIF) {requires  $y \in (0,1)$ }:

$$f(\underline{x}, \underline{y}) = -\frac{1}{M \log(2)} \sum_{m=1}^M [x_m \log(y_m) + (1-x_m) \log(1-y_m)] \quad (\text{Eqn. 8})$$

2. Relative entropy (RENT) {requires  $y > 0$ }:

$$f(\underline{x}, \underline{y}) = -\frac{1}{M \log(2)} \sum_{m=1}^M x_m \log(y_m) \quad (\text{Eqn. 9})$$

3. Euclidean norm (DIST):

$$f(\underline{x}, \underline{y}) = \sqrt{\sum_{m=1}^M (x_m - y_m)^2} \quad (\text{Eqn. 10})$$

4. City block distance (CITY):

$$f(\underline{x}, \underline{y}) = \sum_{m=1}^M |x_m - y_m| \quad (\text{Eqn 11})$$

5

5. Linear correlation coefficient (LCC):

$$f(\underline{x}, \underline{y}) = \frac{\sum_{m=1}^M (x_m - \bar{x})(y_m - \bar{y})}{\sqrt{\sum_{m=1}^M (x_m - \bar{x})^2 \sum_{m=1}^M (y_m - \bar{y})^2}} \quad (\text{Eqn. 12})$$

10 6. Common mean linear correlation coefficient  
(CMLCC):

$$f(\underline{x}, \underline{y}) = \frac{\sum_{m=1}^M (x_m - \bar{x})(y_m - \bar{x})}{\sqrt{\sum_{m=1}^M (x_m - \bar{x})^2 \sum_{m=1}^M (y_m - \bar{x})^2}} \quad (\text{Eqn. 13})$$

7. Root mean power error (RMPE):

$$f(\underline{x}, \underline{y}, \mu) = \sqrt{\frac{1}{M} \sum_{m=1}^M (x_m - y_m)^\mu} \quad (\text{Eqn. 14})$$

15

8. Scaled mean power error (SMPE):

$$f(\underline{x}, \underline{y}, \mu) = \frac{1}{\mu M} \sum_{m=1}^M (x_m - y_m)^\mu \quad (\text{Eqn 15})$$

9. Matrix multiplication (MMLT):

$$f(\underline{x}, \underline{y}) = \sum_{m=1}^M x_m \cdot y_m \quad (\text{Eqn. 16})$$

10-18. Versions of candidate operators 1-9, 'inverted' so that large similarities correspond to small distances and small similarities correspond to large distances, as concisely follows:

$$f(\underline{x}, \underline{y}) = \frac{1}{1 + f'(\underline{x}, \underline{y})} \quad (\text{Eqn. 17})$$

One embodiment of the NSET method and system of this invention can be used to perform process modeling of an instrument calibration verification system as shown in FIGURE 1. The instrument calibration verification system (ICVS) 10 includes Data Server 12, denoted "Data Server or Historical Data," in FIGURE 1, which is used to either obtain real-time samples of the measurements provided by a server program on a plant computer, or to provide previously acquired and stored samples. These measurements are provided to the NSET model 14, which produces the prediction (estimate) of the true process value. Summing module 1b determines the difference (residual) between measurement and model prediction and forwards this result to the sequential probability ratio test (SPRT) module. 16 SPRT is a statistical decision technique that evaluates the difference, or residual, obtained between measurement and model prediction, and determines when drift or failure has occurred. The SPRT has previously been

applied to detection in neural network model based  
ICVSs and in signal monitoring applications  
incorporating other modeling techniques. A full  
description of its role in the ICVS is not provided  
5 here, but it is known to those familiar with the art.  
Lastly, output module 20 provides the results of the  
comparison between measured values and model  
prediction.

An embodiment of the NSET of this invention  
10 applied to an e-commerce setting, and analogous to the  
ICVS embodiment discussed above, is shown in FIGURES 2A  
and 2B. FIGURE 2A illustrates an application of the  
NSET of this invention in an e-commerce setting to  
predict the behavior of a visitor to an e-commerce  
15 site. Database 22 comprises data points representative  
of variables such as gender, socio-economic level,  
geographic location, past purchases, etc. over a large  
population. These data points can span the range of  
the variables they represent. The variables tracked  
20 and stored within database 22 can be arbitrarily  
selected to match the needs of the e-commerce site  
implementing this invention.

When a visitor arrives at an e-commerce site  
implementing this invention, some information is known  
25 about him or her. For example, the visitor's IP  
address can be identified as belonging to a certain  
geographic region. Further information can also be  
known about the visitor from, for example, a sign-in  
sheet at the e-commerce website that requests

demographic information from a customer. Other methods of obtaining information can also be used, such as placing cookies on a visitor's computer or tracking the web page from which a visitor arrived and the path traversed through the e-commerce website. The visitor information that is known can then be compared by the NSET of this invention to statistical historical data contained in database 22.

Database 22 can be comprised of historical information of various parameters as discussed above, and cover a large population base. Database 22 can be initially populated with statistical information that is likely to be relevant, and can then be continuously updated based on new information on visitors to the e-commerce or other relevant information. Database 22 can also be updated with statistical information from other contexts other than e-commerce.

The NSET of this invention will compare a visitor's characteristic values (in the form of data vectors) to the historical data in database 22 to predict or diagnose the visitor's behavior (predict how he or she will behave on the web site) by pulling data points from similar data value patterns as the visitor from database 22. This data selection is done using a clustering technique, as discussed herein.

For example, a visitor's data vector or vectors can be compared to the vectors in a prototype matrix, where the vectors in the prototype matrix comprise statistical characteristic values on height, weight,



zip code, gender, socio-economic status, etc. All of these pieces of information could be contained in historical database 22, comprising statistical characteristic data values obtained over a large population.

When a visitor with known behavior characteristic values (vectors) visits an e-commerce site, the NSET of this invention can, based on those known vectors, pull from historical database 22 data for similar patterns as those of the visitor. For example, if a visitor is known to be a Latin-American female, middle-class, that lives in East Austin, the NSET of this invention can pull from historical database 22 a preset number of patterns that match the visitor's patterns and populate a prototype matrix. The prototype matrix can then be inverted as per the teachings of this invention, and a prediction made for the visitor's behavior based on the statistical data and correlating statistical behavior for that data obtained from database 22. The likely behavior of this visitor can then be predicted as per the teachings of this invention.

The NSET of this invention can adjust its predictions based on new information. For example, if the same visitor as discussed above revisits the same e-commerce location at some future time and divulges, either through direct entry or through some statistical extrapolation (for example the visitor could enter a zip code which the e-commerce site, implementing the NSET of this invention, could recognize as a geographic

area with a large concentration of politicians or a large concentration of wealthy persons, etc.), a new characteristic data value, the NSET can populate a prototype matrix with appropriate data sets more closely matching the user's profile. On this subsequent visit, the NSET of this invention can pull from historical database 22, which can be a pre-organized and indexed database allowing for quick extraction of data, a new set of patterns with which to populate the prototype matrix that incorporates statistical information for wealthy individuals or politicians. Prediction for the behavior of this visitor can then be more accurately obtained. This model fine-tuning can take place several times during the same customer visit, as new information is added to the profile.

The NSET of this invention makes its predictions of the behavior of a visitor (or other measured value) by obtaining a weighted average of the outputs associated with the historical data obtained from database 22. This weighted average is determined based on how close a new data vector is to each of the historical vectors pulled from database 22. If the new vector is identical to an old vector, the weighted average might be weighted to infinity for that vector, while all other historical vectors obtained from database 22 are weighted relatively close to zero. However, if the new vector for a visitor is situated

between two historical vectors, then those two vectors might be weighted equally, and so on.

The NSET of this invention seeks to avoid inverting a large prototype matrix by using clustering techniques to pull only highly relevant data clusters (data points) from historical database 22 to populate the prototype matrix. In this way, whenever a comparison is made between historical data in database 22 and an actual visitor's data vectors, a highly relevant prototype matrix is generated to compare against the visitor. This highly relevant data set is designed to result in a minimum number of data points (the smallest number of data points that will likely yield an accurate prediction) to be inverted in the prototype matrix.

Returning now to FIGURE 2A, database 22 provides the historical statistical data points pulled for a comparison to an actual visitor and provides them to NSET 14. NSET model 14 produces the prediction (estimate) of the behavior of that visitor as discussed previously. Output module 20 then supplies the generated prediction to the e-commerce site implementing this invention. The method of this invention can be implemented as operation instructions executed by processing module 40, and stored within memory 50.

FIGURE 2B illustrates another embodiment of the NSET of this invention used to compare the accuracy of the predictions of NSET module 14 against the actual

behavior of a visitor to an e-commerce site according to the teachings of this invention. Historical database 22 can be updated throughout and following a visit and action by a visitor, and prediction of that visitor's behavior as discussed above in relation to FIGURE 2A. The method and system of this invention can track the visitor's actual behavior following the prediction and subsequently modify the system (similarity function and prototype vectors) to improve future prediction.

The embodiment of this invention shown in FIGURE 2B can be used to compare the actual behavior of a visitor to the predicted behavior as generated by NSET module 14. In a manner similar to that discussed above with respect to FIGURE 1, the actual behavior of the visitor is compared to the visitor's predicted behavior by SPRT module 18, which evaluates the difference (residual) obtained between actual measurement and model prediction.

Using the embodiment of FIGURE 2B, the system and method of this invention can be used to verify the accuracy of the predictions made by the NSET of this invention and to adjust the NSET to increase prediction accuracy. Furthermore, the e-commerce location implementing the method and system of this invention can likewise be adjusted to compensate for the differences between actual or predicted and desired visitor behavior. For example, the goods, services and/or advertising provided by the e-commerce location

can be adjusted to generate more purchases from visitors. Targeted advertising can also be implemented based on the accuracy of the NSET's predictions.

For the purposes of demonstrating system modeling  
5 by the NSET of this invention (as well as  
characterizing the relative performance of the  
candidate operators relative to one another for this  
demonstration application), a sample datafile was  
created containing process measurements from the Oak  
10 Ridge National Laboratory's (ORNL's) High Flux Isotope  
Reactor (HFIR). This datafile spans the research  
reactor's entire fuel cycle, and contains 1636 time-  
measurements. Each time-measurement contains 34 of the  
56 analog signals provided (every 2 seconds) by the  
15 data server on the HFIR's plant computer. The 22  
discarded analog signals include a few process  
variables that never demonstrate any variation from a  
constant measurement value. Other variables, also  
discarded, possessed sufficient physical redundancy  
20 such that inclusion might serve to weaken a  
demonstration of the modeling of an analytically  
redundant modeling system. Some variables represented  
only simple sums or averages of other measured analog  
quantities, and were not included.

25 The constructed datafile contained 1636 data  
vectors representing a stratified sample (every nth  
sample) of an entire fuel cycle (startup transient,  
full power operation, and shutdown transient). This  
data was clustered using a hyperbox clustering

technique, so that a pre-specified number (100 in this example) of data clusters were created. The datapoint closest to the center of each cluster centroid was included in the prototype matrix.

5           For the prototype matrix created, the 18 candidate operators were each explored by computing the sum-squared error (SSE) between NSET model estimation and measurement of the original datafile. As previously mentioned (specifically, in equations 8 and 9), some of  
10   the operators require a particular domain. In addition, demonstration of the effect of scaling upon estimation was desired. To these ends, each operator was first used and evaluated with raw (non-scaled) data. Linear scaling of the data to [0.1, 0.9] was  
15   then performed. The range (the central region of the sigmoid function) of the linearly scaled data was arbitrarily chosen, as previously existing routines for artificial neural network input/output scaling were conveniently available. The estimation was repeated  
20   for each operator, and the estimates were 'de-scaled' for evaluation. Finally, this estimation was performed similarly as before with the linear scaling case, but with mean-centered unit-variance (MCUV) scaling and de-scaling. MCVU scaling results in each individual  
25   scaled time-series possessing a mean of zero and variance of one. The results of this demonstration are given below in Table 1. The columns and rows of Table 1 respectively correspond to the scaling method and the similarity/distance operator utilized.

Table 1. Test Prediction Error Results for  
forward and Inverse Distance/Similarity Operators

	Forward Test SSE			Inverse Test SSE		
	None	Linear	MCUV	None	Linear	MCUV
BERR		2.52E+06			7.43E+03	
RENT		1.10E+07			6.93E+04	
DIST	6.68E+03	1.25E+04	1.45E+04	1.29E+07	4.83E+04	2.60E+05
CITY	4.85E+03	1.53E+04	1.51E+04	4.22E+07	1.54E+05	7.86E+05
LCC	1.27E+06	4.02E+06	6.16E+05	6.99E+04	3.02E+07	7.97E+07
CMLCC	3.09E+03	5.12E+05	1.54E+05	2.64E+03	1.10E+08	3.42E+08
RMPE	6.68E+03	1.25E+04	1.45E+04	7.61E+05	1.63E+04	4.17E+04
SMPE	3.58E-12	1.42E-13	1.72E-15	2.65E+06	3.27E+02	6.48E+03
MMLT	9.38E-10	5.84E-13	2.33E-16	6.33E+04	3.95E+06	2.94E+09

5

Table 2, following, provides the recall estimation rate on a 200 MHz Pentium PC for only the 'forward' version of each operator, as the inverse versions were found to differ by less than 1%. The operators were initially formulated such that all arithmetic was performed as a series of scalar operations upon the elements of two vectors. The 'vectorized' versions are enhanced such that the operators involve a series of vector operations upon matrices, and are thereby significantly improved in MATLAB, the interpreted mathematical programming environment used for NSET development. Future recoding in C++ is expected to bring about further and more significant recall estimation rate improvements.

20

*Table 2. Recall Estimation Rate For Various  
Distance/Similarity Operators*

	Estimations/Second	
	Original	Vectorized
BERR	16.73	45.70
RENT	51.79	87.67
DIST	39.77	142.19
CITY	49.31	123.56
LCC	16.56	55.22
CMLCC	15.21	41.74
RMPE	43.50	90.16
SMPE	46.32	95.29
MMLT		551.61

5

Illustration of inversion difficulties is provided below in Table 3. The last three columns consist of a reciprocal condition number associated with the recognition matrix for each combination of operator and scaling combination. A value near one is associated with an easily invertible matrix, while values approaching zero are associated with near singular matrices. The first three columns of Table 3 correspond to an index created to describe the quality of the pseudo-inverse obtained from the recognition matrix. The index is obtained by pre-multiplying the recognition matrix by its pseudo-inverse. The sum square difference is then computed between the result and the same size identity matrix. Of particular interest is the fact that the result obtained, in all cases, closely approaches an integer. Recognition matrices that possess a reciprocal condition number

10

15

20



very close to zero ( $< 10^{-10}$ ) result in positive integer values (within several decimal places), while greater condition numbers foretell indices near zero. This is consistent with the SVD's method of disregarding more singular values when presented more poorly conditioned matrices.

Table 3. Recognition Matrix Inversion and Invertability Indicators

	Pseudo-inverse Error Index			Reciprocal Condition Number		
	None	Linear	MCUV	None	Linear	MCUV
BERR		65			3.18E-19	
RENT		66			6.94E-20	
DIST	0	0	0	1.60E-04	4.99E-04	4.09E-04
CITY	0	0	0	6.08E-05	1.70E-04	1.79E-04
LCC	67	67	67	9.25E-20	2.59E-19	6.48E-19
CMLCC	5	0	0	2.36E-15	7.48E-10	2.44E-08
RMPE	0	0	0	1.60E-04	4.99E-04	4.09E-04
SMPE	64	64	64	3.71E-21	5.37E-20	3.05E-21
MLT	66	66	66	1.81E-20	1.29E-19	1.15E-20

The root mean power error (RMPE) and scaled mean power error (SMPE) operators contain a variable parameter known as the Minkowski parameter. This parameter influences the proportional effect of each component's difference upon the entire vector distance. At low parameter values, all of the component errors affect the distance. As the Minkowsky parameter is allowed to increase, the distance decreases asymptotically towards the largest component error. As displayed in FIGURE 2, this parameter was varied to

observe the effect upon test prediction error and yielded curve 200.

As sum-squared error is a useful comparison metric between prediction methods, but not a very informative  
5 standalone statistic, FIGURES 3-12 display measured and predicted individual signals over the entire test set - first with no measurement error, then with measurement error, and finally, with the diagnosis of the calibration verification system. FIGURE 3 displays  
10 both the measured coding tower inlet temperature 300 and the estimated cooling tower inlet temperature 310, which overlap significantly as shown. The NSET estimate is obtained using candidate operator #3, the Euclidean distance operator. This operator was found  
15 to provide the best combination of estimation accuracy and insensitivity to measurement error for this example.

FIGURE 4 contains the measured cooling tower inlet temperature 400, this time incorporating a synthetic  
20 linear drift of 0.002% full-scale (%FS) per sample, as well as the 'DIST operator and no scaling version of the NSET estimate 420. Note that the synthetic measurement drift present from sample 751 to sample 1250 is only negligibly present in the NSET estimate  
25 420. This immunity to measurement error makes the NSET an appropriate modeling engine in a calibration monitoring system.

FIGURE 5 redisplay the measurement and estimate in increased detail. FIGURE 6 displays the diagnosis

of an NSET-based ICVS system on the same scale. The output of the SPRT module, contained in curve 600 of FIGURE 6, corresponds to the drift diagnosis. The diagnosis assumes 5 discrete values that correspond to different states of measurement error. Gross drift, indicated by  $\pm 1.0$ , has been arbitrarily defined as 30%FS. Fine drift, indicated by  $\pm 0.5$ , has arbitrarily been set to be 0.5%FS. No drift is indicated by an SPRT output of 0.0. These values can be set differently as desired.

These sensitivity settings generally result in an appropriate alarm when the measurement has drifted halfway to one of the setpoints. An alarm at this halfway point denotes that the SPRT has decided that it is more likely than not that the system has drifted. As illustrated in FIGURE 6, the SPRT catches the drift after 132 samples, corresponding to a drift of 0.26%FS.

One behavior of the SPRT detection system should be noted: after the initial detection, the SPRT redetects the drift with a diminishingly smaller required number of samples. For example, the second detection only takes another 28 samples. Eventually, as the drift approaches 1.0%FS, it is redetected on every other sample. Fewer and fewer samples are required as the drift magnitude gets larger with respect to the nominal variance or random error associated with the residual between the measurement and the NSET estimate.

FIGURES 7-10 display the modeling performance of the NSET using the Euclidean distance operator for the outlet temperature 1 (servo) signal. These FIGURES profile the performance with another typical signal of the same NSET as before, using the same 100 measurement vector prototype matrix, scaling (none), SPRT sensitivities. and the preferred comparison operator.

In FIGURE 7, the measurement and NSET estimate are displayed with no measurement error as curves 710 and 720. FIGURES 8 and 9 demonstrate the behavior of the estimate in the presence of measurement error. A linear drift of 0.002%FS begins at sample 751 and continues until measurement bias of 1%FS has been reached at sample 1250. Measured value curves 810 (FIGURE 8) and 910 (FIGURE 9) and NSET estimate curves 820 (FIGURE 8) and 920 (FIGURE 9) illustrate these findings.

In FIGURE 10, the drift has been detected after 125 samples as shown by SPRT results curve 1010. The magnitude of the drift at this initial point of detection is 0.25%FS. This sensitivity, as previously mentioned, was arbitrarily selected to represent a typical application. Most applications allow for some calibration drift, otherwise continuous calibration is required (drift happens). Much finer sensitivity, when desirable (or even practical) can be obtained by adjusting a single parameter of the SPRT.

As is apparent in Table 1, several operators produced acceptable prediction errors on the natural

test data (containing no drift or measurement error).  
Some of these operators were found to be highly  
susceptible to measurement error and thereby less  
suitable than the chosen DIST/RMPE operator.

5           FIGURES 11 and 12 demonstrate respectively the  
high accuracy clean data estimation and the deficiency  
typical of NSET estimation with some of these  
operators, in this case the common mean linear  
correlation coefficient (CMLCC). First, estimation is  
10 demonstrated with no drift or measurement error in  
curve 1110 of FIGURE 11. Curve 1210 of FIGURE 12 then  
demonstrates the model performance with 0.002%FS drift  
per sample. As seen previously, the CMLCC seems to  
provide very accurate prediction, as long as  
15 measurements are relatively consistent and free of  
measurement error. However, as soon as the measurement  
in FIGURE 12 begin to drift, an undesirable behavior of  
the system incorporating the CMLCC operator comes to  
light: the estimate follows the measurement closely,  
20 even as the measurement drifts! The NSET with the  
CMLCC, as well as with other unsuitable operator  
choices, cannot predict drift and other forms of  
measurement error.

25           A nonlinear state estimation technique (NSET) has  
been developed as a generalization of the multivariate  
state estimation technique (MSET), which is an  
extension of the multiple regression equation. Several  
nonlinear operators have been employed and explored for  
the measurement similarity/distance pattern comparison

implicit in the NSET. Although several operators provided very good 'clean data' model prediction, only a few operators give robust estimation in the presence of measurement errors.

5           The root mean power error (RMPE) operator can provide the most robust prediction, being virtually impervious to measurement error. The RMPE also provided its most accurate prediction (while retaining its robust character) with a Minkowski parameter of  
10   2.0. This parameter value rendered the RMPE equivalent to its more specific (also independently explored), and conveniently coded (such that it is almost 60% faster) version -- Euclidean distance (DIST).

          In this ICVS example, the NSET outfitted with the  
15   RMPE/DIST operator predicted over 140 measurement vectors (each with 34 components) per second in an interpreted environment. Real time or faster than real time prediction and monitoring of thousands of  
20   variables would be possible for a compiled standalone ICVS (with detection of fractional %FS errors) or other process modeling application incorporating NSET on a dedicated inexpensive personal computer or workstation.

          The method and system of this invention use a regularized NSET that utilizes regularization in the  
25   inversion of the prototype matrix. This regularization takes the form of either truncated singular value decomposition (TSVD) or Tikhonov regularization (popularly known in a standard form as ridge regression) in the inversion of troublesome prototype

matrices. This regularization results in improved matrix conditioning, which improves model stability and prediction accuracy. Ridge regression is preferred over TSVD, as the regularization co-efficient is easier  
5 to optimize automatically.

The ability of the method and system of this invention to use a broader spectrum of distance function results in an advantage of allowing the development of more flexible models that may be  
10 optimized to the particular data set or system to be modeled.

The method and system of this invention can be implemented as operational instructions, stored in memory, and executed by a processing module. The  
15 processing module may be a single processing device or a plurality of processing devices. Such a processing device may be a micro-processor, micro-computer, digital signal processor, central processing unit of a computer or workstation, digital circuitry, state  
20 machine, and/or any device that manipulates signals (e.g., analogue and/or digital) based on operational instructions. The memory may be a single memory device or a plurality of memory devices. Such a memory device may be a random access memory, read-only memory, floppy  
25 disk memory, hard drive memory, extended memory, magnetic tape memory, zip drive memory and/or any device that stores digital information. Note that when a processing module implements one or more of its functions, via state machine or logic circuitry, the

memory storing the corresponding operational instructions is embedded within the circuitry comprising the state machine or logic circuitry.

5 The NSET method and system of this invention can be implemented over a computer network, such as the internet, and can obtain data points from a database 22 that need not be co-located with the processing device.

10 Although the present invention has been described in detail herein with reference to the illustrative embodiments, it should be understood that the description is by way of example only and is not to be construed in a limiting sense. It is to be further understood, therefore, that numerous changes in the details of the embodiments of this invention and  
15 additional embodiments of this invention will be apparent to, and may be made by, persons of ordinary skill in the art having reference to this description. It is contemplated that all such changes and additional embodiments are within the spirit and true scope of  
20 this invention as claimed below.



WHAT IS CLAIMED IS:

1. A method for modeling the behavior of a visitor to an e-commerce location, comprising the steps of:

5 obtaining one or more visitor characteristic values;

developing a model of the visitor's behavior according to a nonlinear state estimation technique (NSET); and

10 estimating a set of visitor behavior characteristic values with said model that model said visitor's behavior.

2. The method of Claim 1, further comprising the step of predicting said visitor's future behavior at said e-commerce location based on said set of behavior characteristic values.

3. The method of Claim 1, further comprising the steps of:

determining a set of residuals between said set of estimated behavior characteristic values and a set of actual behavior values; and

statistically monitoring said set of residuals.

25

4. The method of Claim 3, further comprising the step of adjusting said NSET to compensate for said residuals.

5. The method of Claim 3, further comprising the step of adjusting said e-commerce location to compensate for residuals between the desired behavior and the predicted behavior.

5

6. The method of Claim 5, wherein adjusting said e-commerce location comprises adjusting goods, services and/or advertising provided at said e-commerce locations.

10

7. The method of Claim 1, further comprising providing goods, services, and/or advertising at said e-commerce location based on said estimated behavior characteristic values.

15

8. The method of Claim 1, wherein developing a model further comprises the step of selecting a representative sample data set based on said visitor characteristic values from a set of statistical characteristic data within a historical database.

20

9. The method of Claim 8, further comprising the step of populating a prototype matrix with vectors comprising said statistical characteristic data and inverting said prototype matrix according to said NSET.

25

10. The method of Claim 1, wherein said visitor characteristic values comprise visitor demographic information and visitor purchase habits.

11. The method of Claim 1, wherein said nonlinear state estimation technique functions as an regressive model.

5

12. The method of Claim 11, wherein said auto-associative model is not an iterative process and comprises at least one operation selected from the group consisting of single matrix multiplication, inversion and decomposition.

10

13. The method of Claim 1, further comprising the step of:

selectively determining a set of variables for said model; and

15

compactly representing said model as a matrix of observed states of said set of variables for said model.

20

14. The method of Claim 13, wherein selectively determining a set of variables for said model comprises discarding variables that demonstrate an effect on said model below a threshold value.

25

15. The method of Claim 1, further comprising the steps of:

measuring a set of actual behavior values for said visitor based on said visitor's actual behavior at said e-commerce location; and

comparing said set of estimated behavior characteristic values and said set of actual behavior values with at least one similarity operators.

5           16. The method of Claim 15, wherein said similarity operators are selected from the group consisting of:

          Bernoulli difference;  
          Relative entropy;  
10          Euclidean norm;  
          City block distance;  
          Linear correlation coefficient;  
          Common mean linear correlation coefficient;  
          Root mean power error;  
15          Scaled mean power error; and  
          Matrix multiplication.

          17. The method of Claim 16, further comprising the steps of:  
20          selectively determining a set of variables for said model; and  
          compactly representing said model as a matrix of observed states of said set of variables for said model.

25           18. The method of Claim 17, wherein selectively determining a set of variables for said model comprises discarding variables that demonstrate an effect on said model below a threshold value.

19. A system for modeling the behavior of a visitor to an e-commerce location, comprising:

operational instructions for obtaining one or more visitor characteristic values;

5 operational instructions for a NSET for developing a model of the visitor's behavior; and

operational instructions for estimating a set of visitor behavior characteristic values with said model that model said visitor's behavior.

10

20. The system of Claim 19, further comprising operational instructions for predicting said visitor's future behavior at said e-commerce location based on said set of behavior characteristic values.

15

21. The system of Claim 19, further comprising:

operational instructions for determining a set of residuals between said set of estimated behavior characteristic values and a set of actual behavior values; and

20

operational instructions for statistically monitoring said set of residuals.

22. The system of Claim 21, further comprising operational instructions for adjusting said NSET to compensate for said residuals.

25

23. The system of Claim 19, further comprising:

operational instructions for determining a set of residuals between said set of estimated behavior characteristic values and a set of desired behavior values; and

5        operational instructions for statistically monitoring said set of residuals.

24. The system of Claim 23, further comprising the step of adjusting said e-commerce location to  
10        compensate for said residuals.

25. The system of Claim 24, wherein adjusting said e-commerce location comprises adjusting goods, services and/or advertising provided at said e-commerce  
15        locations.

26. The system of Claim 19, further comprising operational instructions for providing goods, services, and/or advertising at said e-commerce location based on  
20        said estimated behavior characteristic values.

27. The system of Claim 19, further comprising a processing module to execute operational instructions and memory operationally coupled to said processing  
25        module for storing data and operational instructions.

28. The system of Claim 19, wherein developing a model further comprises selecting a representative sample data set based on said visitor characteristic

values from a set of statistical characteristic data within a historical database.

29. The system of Claim 28, further comprising  
5 operational instructions for populating a prototype matrix with vectors comprising said statistical characteristic data and inverting said prototype matrix according to said NSET.

10 30. The system of Claim 19, wherein said visitor characteristic values comprise visitor demographic information and visitor purchase habits.

15 31. The system of Claim 19, wherein said nonlinear state estimation technique functions as an regressive model.

20 32. The system of Claim 31, wherein said regressive modeling is not an iterative process and comprises at least one operation selected from the group consisting of single matrix multiplication, inversion and decomposition.

25 33. The system of Claim 19, further comprising:  
operational instructions for selectively determining a set of variables for said model; and  
operational instructions for compactly representing said model as a matrix of observed states of said set of variables for said model.

34. The system of Claim 33, wherein selectively  
determining a set of variables for said model comprises  
discarding variables that demonstrate an effect on said  
5 model below a threshold value.

35. The system of Claim 19, further comprising:  
operational instructions for measuring a set of  
actual behavior values for said visitor based on said  
10 visitor's actual behavior at said e-commerce location;  
and

operational instructions for comparing said set of  
estimated behavior characteristic values and said set  
of actual behavior values with at least one similarity  
15 operators.

36. The system of Claim 35, wherein said  
similarity operators are selected from the group  
consisting of:

20 Bernoulli difference;  
Relative entropy;  
Euclidean norm;  
City block distance;  
Linear correlation coefficient;  
25 Common mean linear correlation coefficient;  
Root mean power error;  
Scaled mean power error; and  
Matrix multiplication.



37. The system of Claim 36, further comprising:  
operational instructions for selectively  
determining a set of variables for said model; and  
operational instructions for compactly  
5 representing said model as a matrix of observed states  
of said set of variables for said model.

38. The system of Claim 37, wherein selectively  
determining a set of variables for said model comprises  
10 discarding variables that demonstrate an effect on said  
model below a threshold value.

39. A method for verifying a calibration status of a set of process sensors, comprising the steps of:  
developing a model of the set of process sensors according to a nonlinear state estimation

5 technique (NSET);

estimating a set of estimated process values for calibrated sensors with said model;

measuring a set of process values with a set of sensors wherein a calibration status of said set of  
10 sensors is unknown;

determining a set of residuals between said set of estimated process values and a set of actual process measurements; and

statistically monitoring said set of residuals.

15

40. The method of Claim 39, wherein said nonlinear state estimation technique functions as an regressive model.

20

41. The method of Claim 40, wherein said auto-associative model is not an iterative process and comprises at least one operation selected from the group consisting of single matrix multiplication, inversion and decomposition.

25

42. The method of Claim 38, further comprising the step of:

selectively determining a set of variables for said model; and

compactly representing said model as a matrix of observed states of the set of process sensors for said set of variables for said model.

- 5           43. The method of Claim 42, wherein selectively determining a set of variables for said model comprises discarding variables that demonstrate an effect on said model below a threshold value.

44. A method for verifying a calibration status of a set of process sensors, comprising the steps of:

developing a model of the set of process sensors according to a nonlinear state estimation technique;

5       estimating a set of estimated process values for calibrated sensors with said model;

measuring a set of process values with the set of sensors wherein a calibration status of the set of sensors is unknown; and

10       comparing said set of estimated process values and said set of actual process measurements with at least one similarity operators.

45. The method of Claim 44, wherein said  
15 similarity operators are selected from the group consisting of:

Bernoulli difference;

Relative entropy;

Euclidean norm;

20       City block distance;

Linear correlation coefficient;

Common mean linear correlation coefficient;

Root mean power error;

Scaled mean power error; and

25       Matrix multiplication.

46. The method of Claim 45, further comprising the steps of:

selectively determining a set of variables for  
said model; and

compactly representing said model as a matrix of  
observed states of the set of process sensors for said  
5 set of variables for said model.

47. The method of Claim 46, wherein selectively  
determining a set of variables for said model comprises  
discarding variables that demonstrate an effect on said  
10 model below a threshold value.

48. An instrument calibration verification system comprising:

a data server to provide real-time measurements or recorded historical measurements of a process sensor;

5 a NSET model that produces an estimate of a true process value;

a statistical decision module that determines and evaluates a difference or residual between said measurements and said estimate of said NSET model in  
10 order to determine when a drift or failure of said process sensor has occurred.

49. The instrument calibration verification system of Claim 48, wherein said statistical decision  
15 module is a SPRT module.

50. The instrument calibration verification system of Claim 48, wherein said statistical decision model compares said measurements and said estimate of  
20 said NEST model with at least one similarity operators.

51. The instrument calibration verification system of Claim 50, wherein said similarity operators are selected from the group consisting of:

25 Bernoulli difference;  
Relative entropy;  
Euclidean norm;  
City block distance;  
Linear correlation coefficient;

Common mean linear correlation coefficient;  
Root mean power error;  
Scaled mean power error; and  
Matrix multiplication.

5

52. The instrument calibration verification system of Claim 48, wherein said NEST model is produced by the steps of:

10 selectively determining a set of variables for said model; and

compactly representing said model as a matrix of observed states of the set of process sensors for said set of variables for said model.

15

53. The instrument calibration verification system of Claim 52, wherein selectively determining a set of variables for said model comprises discarding variables that demonstrate an effect on said model below a threshold value.

20

54. The instrument calibration verification system of Claim 48, further comprising a variable monitoring system to provide real-time measurements or recorded historical measurements of a set of input  
25 variables for said NSET model.

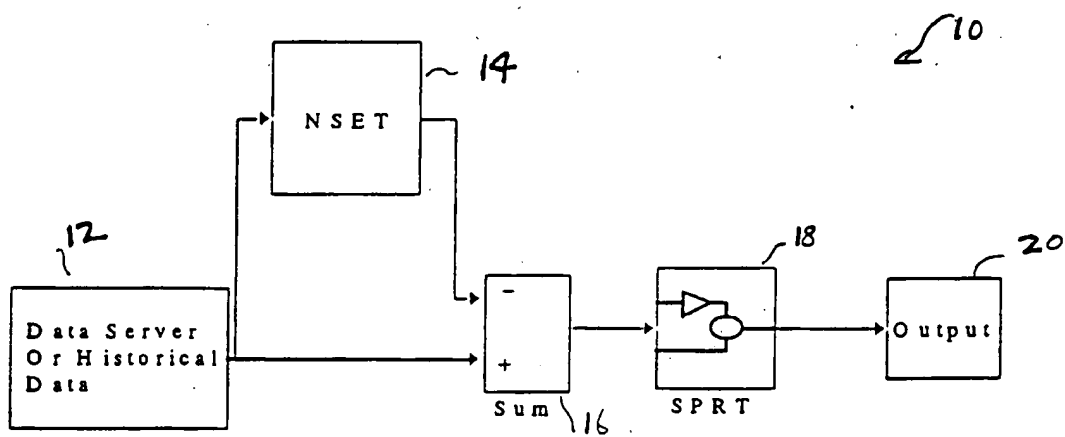


FIG. 1A



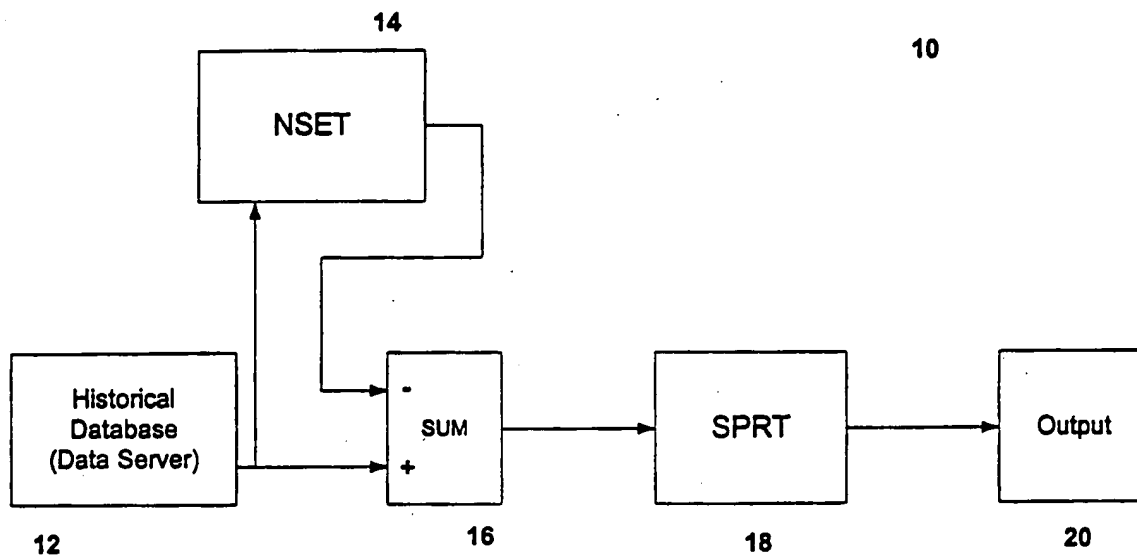


FIGURE 1B

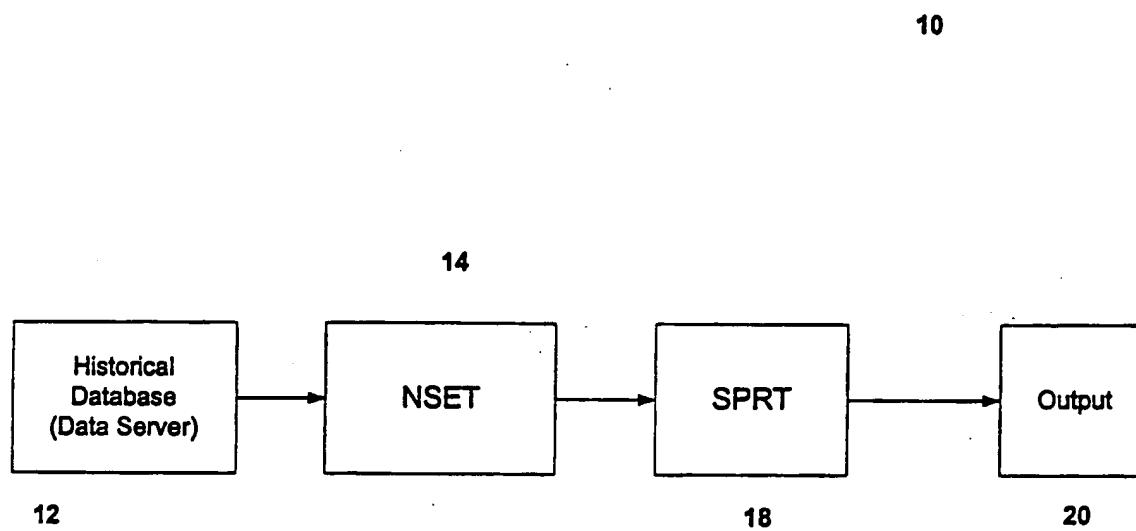


FIGURE 1C

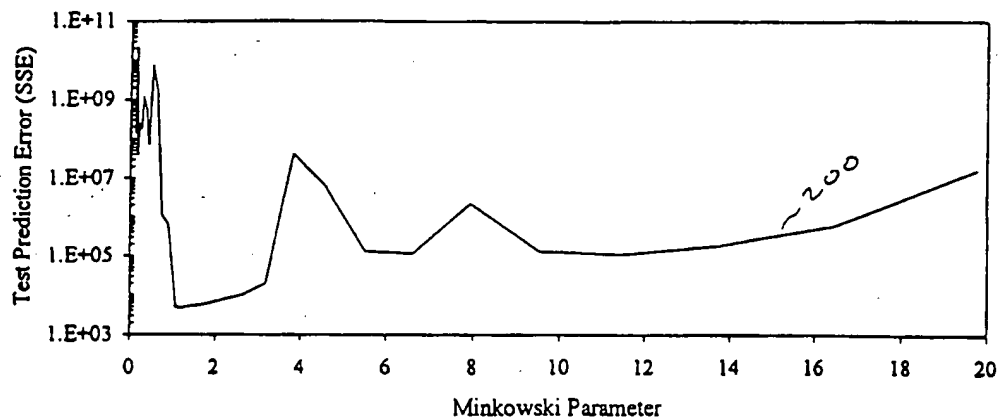


Figure 2. NSET Test Prediction Error Using RMPE Operator and Varying Minkowski Parameter

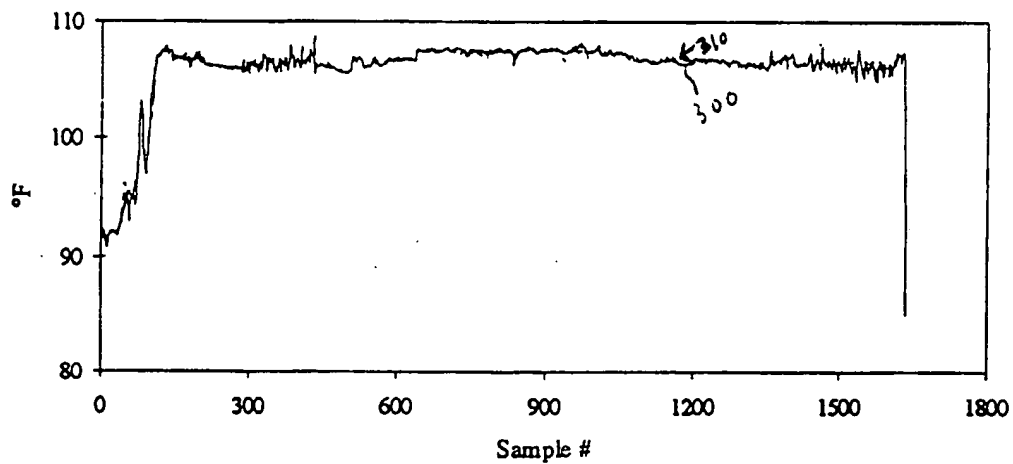


Figure 3. Estimated and Measured Cooling Tower Inlet Temperature (DIST - no Scaling)

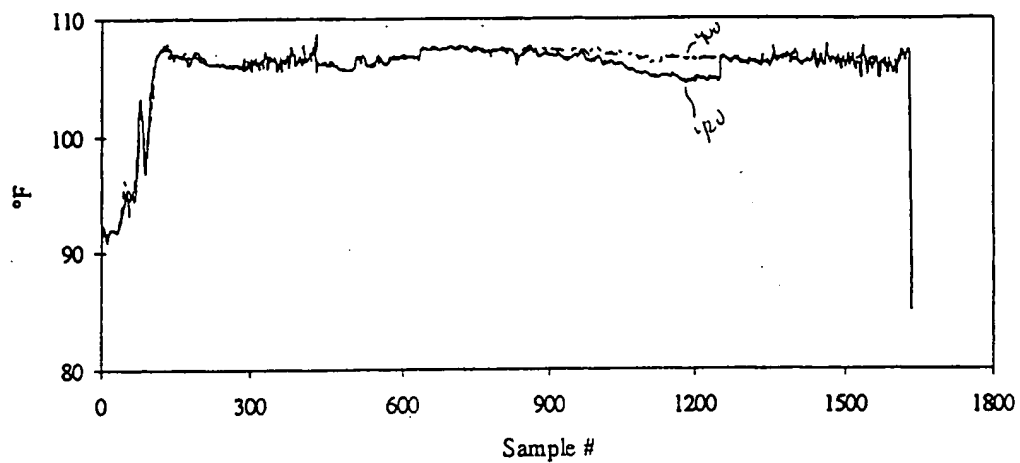


Figure 4. Drifted Measurement and Estimate of Cooling Tower Inlet Temperature (DIST - no scaling)

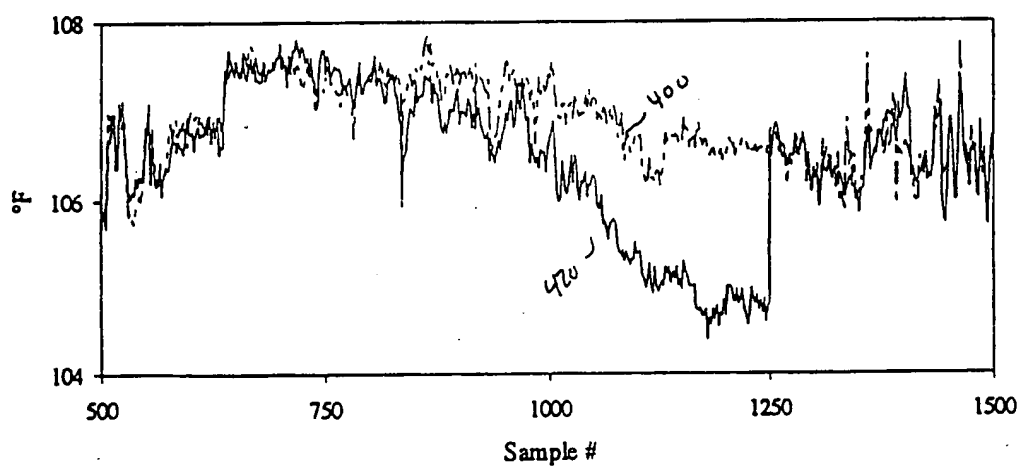


Figure 5. Drifted Measurement and NSET Estimate (Drift commencing at Sample 751)

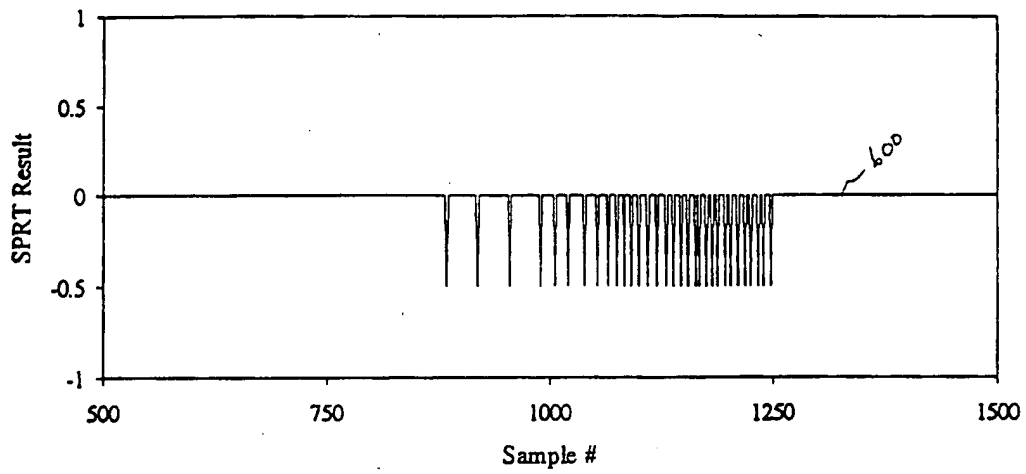


Figure 6. SPRT Drift Diagnosis from NSET (DIST - no scaling) ISCV System

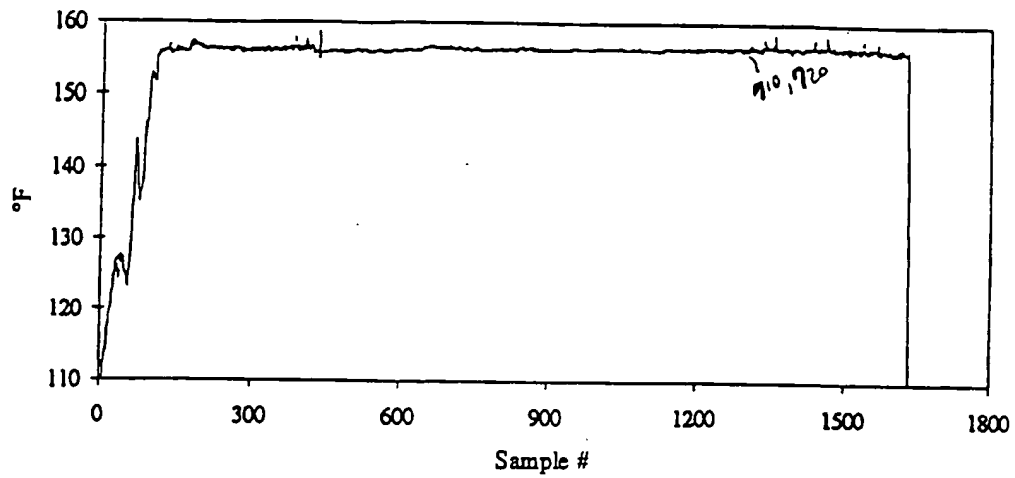


Figure 7. Estimated and Measured Outlet Temperature 1 (Servo) (DIST - no Scaling)

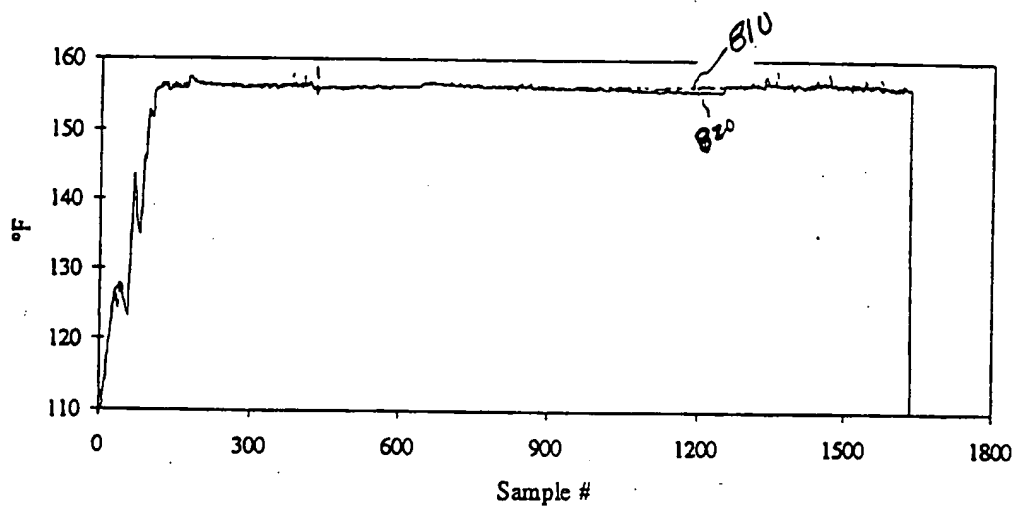


Figure 8. Drifted Measurement and Estimate of Outlet Temperature 1 (Servo) (DIST - no scaling)

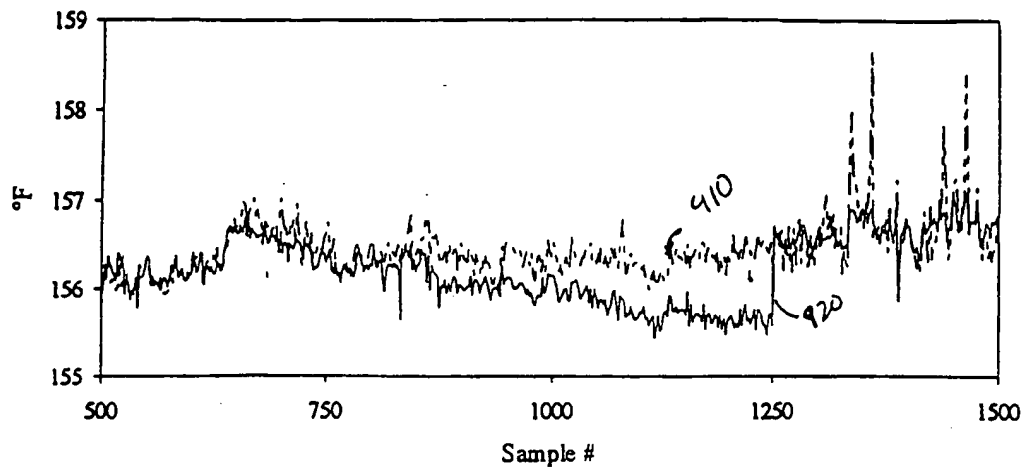


Figure 9. Drifted Measurement and NSET Estimate (Drift commencing at Sample 251)

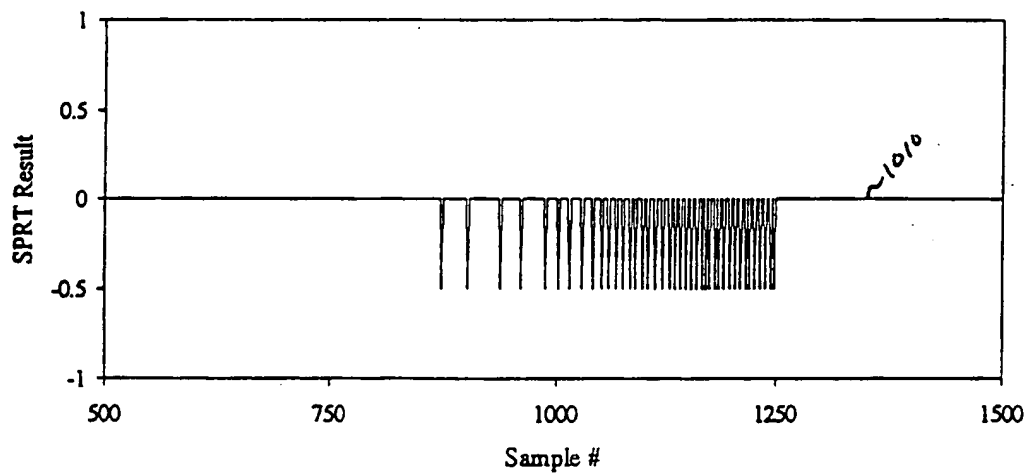


Figure 10. SPRT Drift Diagnosis from NSET (DIST - no scaling) ISCV System

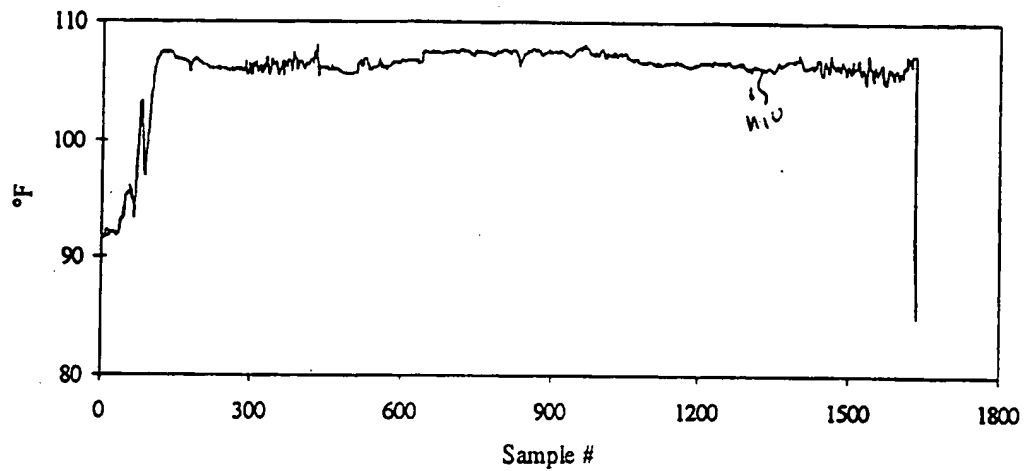


Figure 11. Cooling Tower Inlet Temperature Measurement and CMLCC NSET Estimate

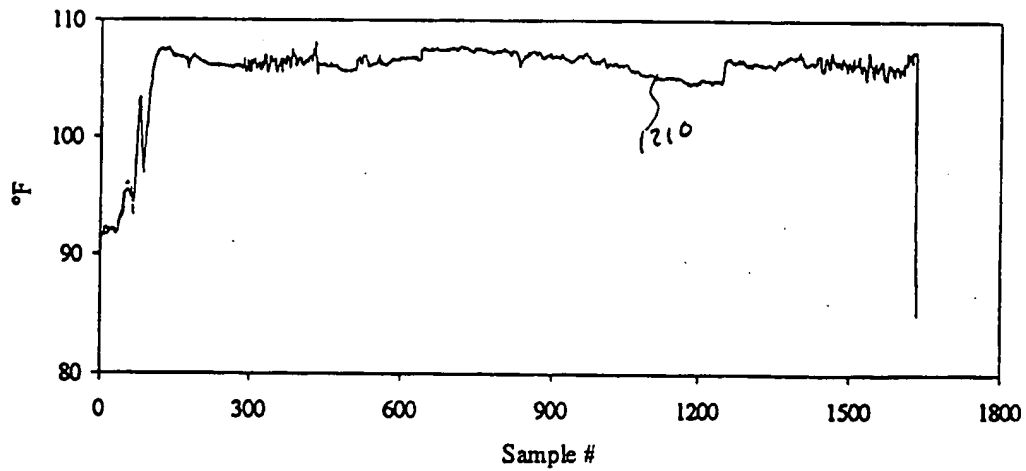


Figure 12. Cooling Tower Inlet Temperature Measurement (Drifted) and CMLCC NSET Estimate